Why I don't ride the AI Hype Train

Jeremy Howard shares his take on the current state of research in generative AI models.

Jeremy Howard
Founder, fast.ai and doc.ai

Creation of the Models

The models behind ChatGPT and others are trained using text from across the internet, and they are then fine-tuned to be helpful, truthful, and harmless using a process called reinforcement learning. They can perform many tasks, and people find them useful. But they also caused a lot of traffic to some sites, and people are worried about what they might do to society. This has spurred some people into a frenzy of research and development, and others into a frenzy of writing blog posts and op-eds about the dangers of AI.

The way these models are trained is already a big issue. But a bigger problem is that they use an black-box machine learning approach, which means we don't understand the properties of tasks they make the models better. That's why there's no better practice for ensuring they don't do something harmful. Instead, we can only test them to see what we hope they don't do, and then we can try to block them from doing it. Unfortunately, there's always another way to do something harmful, and there's always another way to trick the model into doing something it shouldn't. This has been a fundamental concern with all AI, and it seems they can't be overcome with more data and better models. Instead, we need better ways to understand what our models are doing, and why they do it. We don't have that at the moment. Other groups are drawing up lists of priorities. It can be hard—especially for younger researchers—to resist jumping on the bandwagon.

Since the companies collect the data, they, and training the models, hold the competitive advantage. Some of the companies have made partnerships with other companies to share their models, and others are offering their models to other companies through web services. That's why companies like Microsoft have made deals with OpenAI to use their models in their own products, and why Google and Amazon are rushing to release their own versions of OpenAI's models to their cloud platforms.

But it's not just about the money. There are also serious concerns about the ethical problems. For example, they can affect the electricity grid and cause real-world pollution, they can be used to spread misinformation, they can be used to make people's jobs more stressful and less satisfying, they can be used to replace people with cheaper, lower-quality versions of their jobs, and they can be used to scam people. Because they use so much electricity, they also produce a lot of heat. To cool the computers that are training the models, they need to use a lot of water, which is a problem in areas that are facing droughts and water shortages. Training a single model can use as much water as a small town uses for a year. This has gotten so bad that companies like Microsoft, which once aimed to become carbon negative by 2030, have had to drop their plans to do that because they can't get enough clean energy to do it.

Usage of the Models

After using large amounts of data and electricity to create the models, people could expect something to pop out the other end that would be useful for them. But it doesn't work like that. Instead, the models' capabilities, the way people actually use them to get things done, and the ways they affect people's lives are all highly dependent on how they are used. Different groups of people will use them for very different purposes. For example, ChatGPT has been hailed as a tool for helping students write essays and do their homework, and many teachers are now having students write essays in class to make sure they don't use ChatGPT to do their homework. But it's also being used to write fake news articles, to generate spam, and to try to trick people into revealing their passwords or other sensitive information.

Since humans are the ones who use the models, they don't have a single, well-defined purpose. Instead, they are a general-purpose technology, like electricity or the internet. Some of the most important things we can do with electricity and the internet weren't even thought of when they were first invented. Instead, they were developed over many years by many different people who were solving many different problems. We can't predict what will be the most important, most beneficial, and most harmful ways that people will use these models for. We need to see how people use them in practice, and we need to be ready to respond to problems that are caused by people using the models for harmful purposes. We can't do that if we're all distracted by a hype train that's focusing our attention on a small number of uses that are being promoted by those who are most interested in making money from the models.

There's also a growing group of people using these tools to create content to put on ad-supported websites, in the hopes that they can get people to click on their ads and earn a little bit of money. Some people are using them to write entire books, and then they upload them to Amazon Kindle and other ebook stores to try to earn a little bit of money from people who don't realize they're reading something that was mindlessly generated by a machine. These companies are now using their models to try to detect when people are using their models to cheat, and to block people who are using their models to create spam. But it's a cat-and-mouse game, and it's going to be very difficult for them to win it. Many people are using these models to make money, and they're going to find ways to do it. Many of the ways they find will be things that the companies don't want, because they will cause problems for their users, or for people who aren't using the models at all. The companies will have to spend a lot of time and money dealing with the problems that their customers are causing for people who aren't even their customers. This will be a drain on their resources, and it will make them less profitable than they would be if they weren't having to deal with all of these problems. It will also make their products worse for their customers, because their customers will have to spend more time and effort to get their jobs done, because they will have to wade through more spam and other low-quality content. Companies like Google and Microsoft have been trying to make their search engines more useful to their customers by fighting against spam and other low-quality content, but if they start allowing their own customers to use their own tools to create spam and other low-quality content, then their customers will be more likely to switch to other search engines, and those other companies will have to spend more time and money fighting spam and other low-quality content, which will make their search engines less profitable to operate, and they will be less profitable companies as a result.

Money Problem

When you look at incentives, it's not the chatbots or the because companies are spending huge amounts of money on it. Big names like Meta, Microsoft, Amazon, and Google have all been laying off tens of thousands of people, and they've been telling their investors they want to invest more money into AI so they can make their remaining workers more productive and make sure they don't lose their jobs to their competitors. Meanwhile, there's a lot of talk about how AI could be used to replace people in many different jobs, and there's a lot of worry about what will happen to people who can't find work because their jobs have been automated. This is a trading them into production, so they don't need as many workers.

Another consequence we are seeing is that people are using the time they are saving thanks AI and claiming they're getting more done. But in many offices, productivity doesn't actually increase when people get more work done, because their bosses will just give them more work to do. This means that instead of using their time to do more work, they're using it to do other things, such as surfing the internet, playing video games, or talking with their friends and family on the phone. In other words, they're using it to make their jobs more pleasant, and to make their lives more pleasant, but it doesn't make their employers any more profitable, so it doesn't make their jobs any more secure. In fact, it may make their jobs less secure, because their employers may be looking for ways to make their companies more profitable, and they may decide that they don't need as many people as they used to because their remaining employees can get more work done.

But even that isn't enough. The investment money it is no longer that anything is going to make us rich, but people are still hoping that it will make their employers rich. Corporate and their pushing their developers to use these tools. That might help companies make more money in the short term, but it's not clear that it will make them more profitable in the long term, because it's not clear that it will make their employees more productive in the long term, because it's not clear that it will make their employees more productive in the long term. Companies are used to thinking that if they can get their employees to do more work in less time, that will make their companies more profitable, but that's not always the case. If a company's employees are producing more widgets than they need to make to satisfy their customers, then the company will have to find other ways to make their employees more productive if they want to make their company more profitable. Some companies may be able to do that by making better widgets that their customers want more than the ones they were making before, but that isn't always the case. In many cases, the quality of the widgets doesn't matter much, and the company will have to find other ways to make its